

REMARKS

Claims 1-18 are pending in the application. Claims 1, 6, 7, 10, 16 and 17 have been amended. Claims 19 and 20 are added. Support for these claims are found, for example, at page 16, line 14 to page 19, line 16 of the specification as originally filed. No new matter is introduced.

Applicants thank the Examiner for taking the time to discuss the Office Action in a telephonic interview on February 8, 2006. The Examiner's input is much appreciated.

I. Amended Claims

Base Claims 1 and 10 have been amended to recite "...at least one non-binary tests and at least one test: (a) examining syntax or grammar; (b) determining whether a predefined piece of data or keyword appears in URLs in the subject Web page or in URLs that refer to the subject Web page...." Support for this amendment is found at, for example, Claim 7 as originally filed.

Base Claims 1 and 10 have been further amended to recite "...or (c) examining page format or style order other than the position of data or a keyword in the subject Web page." Support for this amendment is found at, for example, page 17, lines 12-15 of the specification as originally filed.

Dependent Claims 6, 7, 16 and 17 have been amended to delete the limitations which have been merged into respective base Claims 1 and 10. Applicants respectfully request that this amendment be entered.

No new matter has been introduced by way of this amendment.

II. Rejections under 35 U.S.C. § 103(a)

The present invention is directed to computer-implemented methods of determining the content type of a subject Web page. Particularly, in base Claims 1 and 10, the method includes the step of preparing a distinguishing series of tests. In addition to non-binary tests, the distinguishing series of tests include at least one test: (a) examining syntax or grammar; (b) determining whether a predefined piece of data or keyword appears in URLs in the subject Web page or in URLs that refer to the subject Web page; or (c) examining page format or style order

other than the position of data or a keyword in the subject Web page. Examining syntax, grammar, page format or style ascertain, for example, the number of passive sentences, number of sentences without a verb, percentage of verbs in past tenses, number of fonts used, existence of certain characters in determining the content type of the subject Web page. (*See* page 7, lines 8-11 and page 19, lines 4-12 of the present application.) Furthermore, apart from examining the content of the subject Web page, the present invention also inspects specifically the data or keyword appearing in URLs of the subject Web page or in URLs of the neighboring pages that refer to the subject Web page. (*See* page 19, lines 13-16 and page 23, line 16 to page 24, line 25 of the present application.) As explained below, Applicants respectfully request that the Examiner note that these tests involving syntax, grammar, style or the URLs associated with the subject Web page are what distinguish the present invention from the prior art cited by the Examiner.

A. 35 U.S.C. § 103(a) Rejection of Claims 1-3, 6-9, 10-12 and 15-18

Claims 1-3, 6-9, 10-12 and 15-18 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Russell-Falla et al. (U.S. Patent No. 6,675,162) (hereinafter “Russell-Falla”) in view of Chakrabarti et al. (U.S. Patent No. 6,389,436) (hereinafter “Chakrabarti”).

As the Examiner noted, “Russell-Falla does not teach wherein the distinguishing series of test includes at least one or more non-binary tests.” (Office Action, at page 4.) Furthermore, Chakrabarti does not teach, suggest or otherwise make obvious a distinguishing series of tests that involves syntax, grammar or style as now claimed in Claims 1 and 10. Chakrabarti is directed to a hypertext classifier that classifies a new document containing citations (i.e. hyperlinks) to and from other neighboring documents. (*See* col. 26, lines 46-47 of Chakrabarti.) The hypertext classifier classifies a new document based on its text as well as the text in the other neighboring documents. Chakrabarti categorizes a Web page by detecting certain “terms,” which he refers to as “class names, class identifiers, specially tagged terms, or any token in general.” (*See* col. 10, lines 57-64 of Chakrabarti.)

Contrary to the present invention, Chakrabarti would not classify according to the linguistic, stylistic and formatic appearance of the Web page, distinguishing, for example, according to: “[t]he percentage of header lines in the page [that] is in one of the following

ranges”; “[t]he average sentence length in the page [that] is in one of the following ranges; or “[whether] the first sentence of the first paragraph [of the subject Web page] has a date.” (See page 17, lines 9-10 and page 18, lines 17-23 of the present application.)

Furthermore, Chakrabarti examines the URLs written *on* the Web page but not the actual URL *of* the Web page. The classifying scheme of Chakrabarti is analogous to a situation where a possessor of numerous letters is classifying the letters into different categories (i.e. personal, business, or others) based on only the content of a letter but not on the addresses of a sendee and a sender. Chakrabarti may use the hypertext or URLs *on* the Web page and *on* the neighboring page(s) to point the classifier where the test should be conducted. The classifier of Chakrabarti, however, does not search for terms in the character strings *of* the URLs of the current Web page.

Therefore, Chakrabarti does not teach, suggest or otherwise make obvious the distinguishing series of tests involving one or more tests involving syntax, grammar, page style or URLs associated with the subject Web page as now recited in base Claims 1 and 10.

Accordingly, Applicants respectfully request the § 103(a) rejections of base Claims 1 and 10 be withdrawn.

In addition, Claims 2-9, which depend from Claim 1, and Claims 11-18, which depend from Claim 10, are patentable for the same reason as Claims 1 and 10. In particular, Claim 6 is patentable in view of Russell-Falla and Chakrabarti because neither teaches, suggests or otherwise makes obvious a predefined set of tests that includes a potential content type of *press release*. Therefore, in addition to the rationales discussed above, Applicants respectfully request the § 103(a) rejection of Claim 6 be withdrawn because of this independent reason.

B. 35 U.S.C. § 103(a) Rejection of Claims 4, 5, 13 and 14

Claims 4, 5, 13 and 14 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Russell-Falla in view of Chakrabarti in further view of Haug et al. (U.S. Patent No. 6,556,964) (hereinafter “Haug”). Haug is directed to a probabilistic model for determining the meaning of sentences or phrases in medical reports. The Haug model extracts and encodes medical concepts using a Bayesian network.

Claims 4, 5, 13 and 14 depend from base Claims 1 or 10. Therefore, Claims 4, 5, 13 and 14 also include the element of the distinguishing series of tests having both binary and non-

binary tests. As explained above, neither Russell-Falla nor Chakrabarti teach, suggest or otherwise make obvious the distinguishing series of tests having one or more tests involving syntax, grammar or page style of Claims 4, 5, 13 and 14. Furthermore, Haug does not cure this deficiency to make Claims 4, 5, 13 and 14 obvious. Therefore, Applicants respectfully request the § 103(a) rejection of Claims 4, 5, 13 and 14 be withdrawn.

III. New Claims 19 and 20

New Claims 19 and 20 depend from base Claims 1 and 10, respectively. Thus, the foregoing patentable distinctions apply to Claims 19 and 20. Acceptance is respectfully requested.

CONCLUSION

In view of the above amendments and remarks, it is believed that all claims (claims 1-20) are in condition for allowance, and it is respectfully requested that the application be passed to issue. If the Examiner feels that a telephone conference would expedite prosecution of this case, the Examiner is invited to call the undersigned.

Respectfully submitted,

HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

By _____
H. Joon Chung
Registration No. 52,748
Telephone: (978) 341-0036
Facsimile: (978) 341-0136

Concord, MA 01742-9133

Dated:

2/24/08